Preprint
UCRL-JC-141742

Multigrid Algorithms For Solving The Linear Boltzmann Equation Using First-Order System Least Squares Finite Element Methods

B. Chang, B. Lee

This article was submitted to
Nuclear Explosive Code Development Conference
Oakland, CA
October 27, 2000

U.S. Department of Energy
Lawrence
Livermore
National
Laboratory

December 13, 2000

Approved for public release; further dissemination unlimited

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint is made available with the understanding that it will not be cited or reproduced without the permission of the author.

This report has been reproduced
directly from the best available copy.

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information
P.O. Box 62, Oak Ridge, TN 37831
Prices available from (423) 576-8401
http://apollo.osti.gov/bridge/

Available to the public from the
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd.,
Springfield, VA 22161
http://www.ntis.gov/

OR

Lawrence Livermore National Laboratory
Technical Information Department's Digital Library
http://www.llnl.gov/tid/Library.html

Multigrid Algorithms for Solving the Linear Boltzmann Equation using First-Order System Least-Squares Finite Element Methods

B. Chang and B. Lee

1 Introduction

Solving the linear Boltzmann equation in neutron scattering phenomena presents many challenges to standard numerical schemes in computational physics. For an SN discretization, the so-called ray effects pollute the numerical solution. This pollution can be viewed mathematically as "contamination" from a poorly chosen approximating basis set for the angle component of the discretization- i.e., collocation in angle is equivalent to discretization with delta basis functions, which form a poor approximating basis set. Fortunately, a PN discretization, which uses a better approximating basis set (i.e., spherical harmonics), eliminates these ray effects. Unfortunately, solving for the moments or PN equations is difficult. Moments couple strongly with each other, creating a strongly coupled system of partial differential equations (pde's); numerical algorithms for solving such strongly coupled systems are difficult to develop. In this paper, novel algorithms for solving this coupled system are presented. In particular, algorithms for solving the PN discretization of the linear Boltzmann equation using a first-order system least-squares (FOSLS) methodology (c.f. [1]) are presented.

This paper is an extension of the research reported in [2]. In that paper, a preconditioned conjugate gradient iteration with a block diagonal preconditioner was used to solve the system of PN equations. Each block of this preconditioner described only a single diagonal $lm-$to$-lm$ moment coupling, but defined over the whole spatial domain. Thus, successively inverting each block of this preconditioner corresponds to successively solving only the $lm-lm$ equations over the whole spatial domain. However, the numerical results presented in that paper demonstrate the non-scalibility of this algorithm with respect to both the number of moments and the number of spatial nodes used in the PN discretization. This non-scalibility reflects this scheme's inability to handle the strong moment coupling.

In this paper, several algorithms that ameloriate some of the moment coupling are presented. One algorithm consists of a multigrid scheme for the spatial

coupling of the PN discretization. Here, the unknowns are updated momentwise first and then spatialwise (i.e., the running indices for the unknowns are faster for the moment indices than for the spatial indices). In this way, for a Gauss-Seidel relaxation scheme, at each spatial node in turn, every moment is updated before going to the next spatial node so that the full moment coupling are considered at a fixed node. Physically, local conservation is somewhat enforced at each spatial node.

A second algorithm presented in this paper is a preconditioned conjugate gradient iteration with a block diagonal preconditioner that describes the full intra-moment coupling. Each diagonal block describes the full $l - l$ moment coupling (i.e., moments $lm - lm'$ with $-l \leq m, m' \leq l$) over the whole spatial domain. Each of these blocks is solved using a few cycles of the above multigrid scheme restricted to the $l - l$ moment block. Thus, local conservation is not as strongly enforced as the above multigrid scheme, but still more than what is enforced in the method of [2]. Comparing the results of this preconditioned conjugate gradient scheme with the above multigrid scheme will expose the relative strength of the inter- and intra-moment coupling in the PN equations.

This paper proceeds as follows. In section 2, a summary of the FOSLS theory developed in [1] for the isotropic, constant coefficient linear Boltzmann equation is reviewed. This theory shows that locally away from the material interfaces, by appropriately scaling the system of pde's, the second-order moment coupling essentially describes the whole coupled system of pde's. This fact will be used to develop our numerical schemes. In section 3, the PN-h finite element discretization for the FOSLS formulation is developed. The system of pde's is explicitly described, and from this description, it will be shown why some scattering parameter regimes require special discretization and non-standard multigrid schemes, topics which will not be examined in this paper. In sections 4 and 5, respectively, the multigrid and preconditioned conjugate gradient schemes are described. Multigrid components (relaxation and coarse grid correction), methods of homogenization of the fine grid material and scaling coefficients, and parallelization issues will be examined. In section 6, computational scaling studies for homogeneous material problems will be presented for both the multigrid and the preconditioned conjugate gradient schemes. For the multigrid scheme, these results demonstrate scalibility with respect to the number of spatial nodes and mild scalibility degradation with respect to the number of moments. For the preconditioned conjugate gradient scheme, these results demonstrate mild non-scalibility with respect to the number of spatial nodes and to the number of moments. This difference in these two schemes indicate a spatially smooth inter-moment coupling error mode that is not handled by this latter scheme. Also presented in this section are results for the Kobayashi test suite problems (c.f. koba). These numerical results show that material inhomogenities have only little affect on the multigrid convergence rates. Finally, section 7 describes some future algorithmic research.

2 Theory

Let $R \times S^2$ be the product space of a bounded domain $R \subset \Re^3$ of unit diameter with the unit sphere S^2. The single group, time-independent form of the linear Boltzmann equation is

$$\begin{aligned} [\Omega \cdot \nabla + \sigma_t I - \sigma_s P]\psi(\mathbf{x}, \Omega) &= q \qquad (\mathbf{x}, \Omega) \in R \times S^2 \\ \psi(\mathbf{x}, \Omega) &= g \qquad \mathbf{x} \in \partial R,\ \mathbf{n} \cdot \Omega < 0. \end{aligned} \tag{1}$$

Here, σ_t and σ_s are respectively the total cross section and scattering cross section, P is the "scattering" operator

$$[P\psi](\mathbf{x}) := \int_{S^2} \psi(\mathbf{x}, \Omega)\, d\Omega$$

with the normalized differential element

$$d\Omega = \frac{\sin(\theta)\, d\theta d\phi}{4\pi},$$

and $\psi(\mathbf{x}, \Omega)$ is the angular flux of neutrons. We have assumed an isotropic medium.

In the FOSLS formulation of (1), the Boltzmann operator is rewritten as

$$\begin{aligned} \mathcal{L}: &= \Omega \cdot \nabla + \sigma_t(I - P) + \sigma_a P, \\ &= \Omega \cdot \nabla + T \end{aligned} \tag{2}$$

where $\sigma_a = \sigma_t - \sigma_s$ is the absorption cross section and $T := \sigma_t(I - P) + \sigma_a P$. Introducing the scaling operator

$$S = \begin{cases} I & \sigma_t \leq 1 \\ \sigma_t(I - P) + \sigma_a P & \sigma_t \geq 1 \text{ and } \sigma_a \geq \frac{1}{\sigma_t} \\ \sigma_t(I - P) + \frac{1}{\sigma_t} P & \sigma_t \geq 1 \text{ and } \sigma_a \leq \frac{1}{\sigma_t} \end{cases}$$

with inverse

$$\begin{aligned} S^{-1} &= \begin{cases} (I - P) + P & \sigma_t \leq 1 \\ \frac{1}{\sigma_t}(I - P) + \frac{1}{\sigma_a} P & \sigma_t \geq 1 \text{ and } \sigma_a \geq \frac{1}{\sigma_t} \\ \frac{1}{\sigma_t}(I - P) + \sigma_t P & \sigma_t \geq 1 \text{ and } \sigma_a \leq \frac{1}{\sigma_t} \end{cases} \\ &= c_1(I - P) + c_2 P, \end{aligned} \tag{3}$$

the space-angle FOSLS formulation is to minimize the scaled least-squares functional

$$\mathcal{F}(\psi; q, g) := \left\| S^{-\frac{1}{2}}(\mathcal{L}\psi - q) \right\|^2 + 2 \int_{\partial R} \int_{\mathbf{n} \cdot \Omega < 0} (\psi - g)\overline{(\psi - g)} |\mathbf{n} \cdot \Omega|\, d\Omega d\sigma$$

over an appropriate Sobolev space. Note that because of the boundary integral in the least-squares functional, the inflow boundary condition need not be enforced on this Sobolev space. Note also that $\|\cdot\|$ is the L^2 norm over both angle and space.

The appropriate Sobolev space is

$$V := \left\{ v \in L^2(R \times S^2) : \left(S^{-1}\Omega \cdot \nabla v, \Omega \cdot \nabla v\right) + (Tv, v) < \infty \right\}$$

with norm

$$\|v\|_V^2 = \left(S^{-1}\Omega \cdot \nabla v, \Omega \cdot \nabla v\right) + (Tv, v).$$

It was shown in [1] that $\mathcal{F}$ is equivalent to the norm

$$\|v\|_{V_1}^2 := \|v\|_V^2 + \int_{\partial R} \int_{\mathbf{n}\cdot\Omega<0} \psi\overline{\psi} |\mathbf{n} \cdot \Omega| \, d\Omega d\sigma$$

over space V. Thus, functional

$$\begin{aligned} \mathcal{F}(\psi; 0, 0) &= \left(S^{-1}\Omega \cdot \nabla\psi, \Omega \cdot \nabla\psi\right) + \left(S^{-1}T\psi, \Omega \cdot \nabla\psi\right) + \left(S^{-1}\Omega \cdot \nabla\psi, T\psi\right) \\ &\quad + (T\psi, T\psi) + 2\int_{\partial R} \int_{\mathbf{n}\cdot\Omega<0} \psi\overline{\psi} |\mathbf{n} \cdot \Omega| \, d\Omega d\sigma \end{aligned}$$

is equivalent to

$$\|\psi\|_{V_1}^2 = \left(S^{-1}\Omega \cdot \nabla\psi, \Omega \cdot \nabla\psi\right) + (T\psi, \psi) + \int_{\partial R} \int_{\mathbf{n}\cdot\Omega<0} \psi\overline{\psi} |\mathbf{n} \cdot \Omega| \, d\Omega d\sigma$$

That is, the first-order terms $\left(S^{-1}T\psi, \Omega \cdot \nabla\psi\right)$ and $\left(S^{-1}\Omega \cdot \nabla\psi, T\psi\right)$ are majorized by the second-order term $\left(S^{-1}\Omega \cdot \nabla\psi, \Omega \cdot \nabla\psi\right)$.

Now minimizing $\mathcal{F}$ over V is equivalent to solving the variational equation

$$\begin{aligned} a(\psi, w) : &= \left(S^{-1}\mathcal{L}\psi, \mathcal{L}w\right) + 2\int_{\partial R} \int_{\mathbf{n}\cdot\Omega<0} \psi\overline{w} |\mathbf{n} \cdot \Omega| \, d\Omega d\sigma \\ &= (q, S^{-1}\mathcal{L}w) + 2\int_{\partial R} \int_{\mathbf{n}\cdot\Omega<0} g\overline{w} |\mathbf{n} \cdot \Omega| \, d\Omega d\sigma \end{aligned}$$

for all $v \in V$. Because of the norm equivalence, one essentially needs to develop an effective solver or preconditioner for the discrete system corresponding to the bilinear form

$$b(v, w) : = \left(S^{-1}\Omega \cdot \nabla v, \Omega \cdot \nabla w\right) + (Tv, w) + \int_{\partial R} \int_{\mathbf{n}\cdot\Omega<0} v\overline{w} |\mathbf{n} \cdot \Omega| \, d\Omega d\sigma.$$

A scalable solution method for the minimization of the least-squares functional will require a scalable solver for this latter system.

3 Spherical Harmonic-h (PN-h) Finite Element Discretization

One of the advantages of a FOSLS formulation is that it leads to symmetric and positive-definite linear systems. For the Boltzmann equation, this allows such efficient linear system solvers as multigrid and preconditioned conjugate gradient schemes to be used on the PN discretization of the FOSLS variational form. Indeed, standard Galerkin PN discretizations of the Boltzmann equation lead to non-symmetric linear systems that are difficult to solve efficiently. There are efficient Petrov-Galerkin formulations of the SN discretization, but this discretization suffers from the ray effect in the thick and thin regions. Nevertheless, the PN FOSLS method is not immuned from problems itself, as will be shown later.

The PN discretization consists of taking a truncated spherical harmonic expansion of the angular flux:

$$\begin{aligned} \psi(\mathbf{x},\Omega) &\approx \psi_N(\mathbf{x},\Omega) \\ &= \sum_{l=0}^{N}\sum_{m=-l}^{l} \phi_{lm}(\mathbf{x})Y_{lm}(\Omega). \end{aligned} \tag{4}$$

The ϕ_{lm}'s are the moments or generalized Fourier coefficients and the Y_{lm}'s are the spherical harmonics. Substituting ψ_N into bilinear form $a(\cdot,\cdot)$, and testing it against $v(\mathbf{x})Y_{l'm'}(\Omega)$, $l' = 0,...,N$ and $m' = -l',...,l'$, a semi-discretization is obtained. Now because

$$[I-P]\phi_{l,m}(\mathbf{x})Y_{lm}(\Omega) = \phi_{l,m}(\mathbf{x})[(I-P)Y_{lm}(\Omega)],$$

T and S^{-1} simply projects the zero and non-zero moments differently. Moreover, because of the norm equivalence, to analyze this semi-discrete system, only the zeroth-order and second-order terms need to be consider.

For the zeroth-order term, we have

$$(T\psi_N, vY_{l',m'}) = \sum_{lm} < Y_{lm}|T|Y_{lm} > (\phi_{lm}, v)_R, \tag{5}$$

where $< \cdot|A|\cdot >$ is the bra-ket notation for the angular inner product with operator A acting on ket $|\cdot >$, and where $(\cdot,\cdot)_R$ is the spatial inner product. For the second-order term, we have

$$\begin{aligned} &\left(S^{-1}\Omega\cdot\nabla\psi_N, \Omega\cdot\nabla vY_{l',m'}\right) \\ &\quad= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{lm} \left(S^{-1}\Omega_i Y_{lm}\phi_{lm,i}, \Omega_j Y_{l'm'}v_j\right) \\ &\quad= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{lm} < Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} > (\phi_{lm,i}, v_j)_R. \end{aligned} \tag{6}$$

Here, i and j denote spatial differentiation. Note that the sparsity pattern of the second-order term depends on both the spatial differentiation operators and the moment coupling created through

$$< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} > .$$

Consider the diagonal $lm-lm$ element of $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$. This element can be viewed as a full 3×3 tensor describing the "diffusion" interaction of moment ϕ_{lm} with itself. Viewed this way, $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$ is a $(N+1)^2 \times (N+1)^2$ block matrix of 3×3 tensors with each $lm-l'm'$ tensor describing not only the moment coupling of ϕ_{lm} to $\phi_{l'm'}$, but also the spatial anisotropy of this moment coupling. Fortunately, this block matrix of tensor has some structure. To see this, the completeness property

$$\sum_{l''m''} |Y_{l''m''} >< Y_{l''m''}| = I$$

of the spherical harmonics (c.f. Sakurai) is needed. Applying this identity twice, we have

$$\begin{aligned} < Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} > &= \sum_{l''m''} < Y_{lm}|\Omega_i S^{-1}|Y_{l''m''} >< Y_{l''m''}|\Omega_j|Y_{l'm'} > \\ &= \sum_{l''m''} \sum_{l'''m'''} < Y_{lm}|\Omega_i|Y_{l'''m'''} > \\ &\qquad < Y_{l'''m'''}|S^{-1}|Y_{l''m''} >< Y_{l''m''}|\Omega_i|Y_{l'm'} > . \end{aligned}$$

But S^{-1} simply scales the ket $|Y_{l''m''} >$ by

$$\begin{cases} c_1 & l'' \neq 0 \\ c_2 & l'' = 0 \end{cases}$$

(c.f., equation (3)). The orthogonality of spherical harmonics then implies

$$< Y_{l'''m'''}|S^{-1}|Y_{l''m''} >= \begin{cases} \delta_{l''l'''}\delta_{m''m'''}\, c_1 & l'' \neq 0 \\ \delta_{l''0}\delta_{m''0}\, c_2 & l'' = 0, \end{cases}$$

and so,

$$\begin{aligned} < Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} > &= \sum_{l''m''} c_{l''m''} < Y_{lm}|\Omega_i|Y_{l''m''} > \\ &\qquad < Y_{l''m''}|\Omega_j|Y_{l'm'} >, \end{aligned} \tag{7}$$

where $c_{l''m''} = c_1$ if $l \neq 0$ and $c_{00} = c_2$. Moreover, it can be shown that

$$< Y_{\hat{l}\hat{m}}|\Omega_i|Y_{\hat{l}\hat{m}} >\neq 0$$

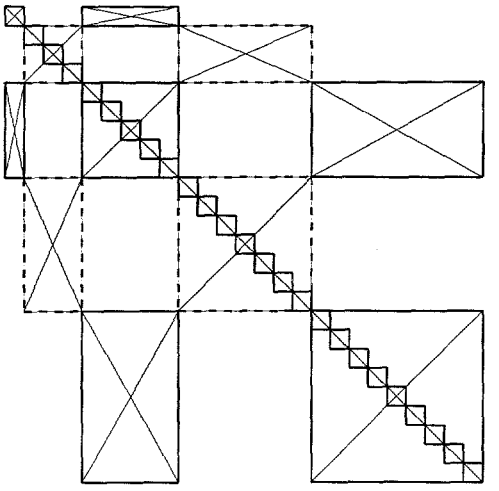

Figure 1: Intra-moment coupling structure with diagonal $lm - lm$ blocks.

only when $\hat{l} = \tilde{l} \pm 1$. Thus, $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$ is the weighted product of two block tridiagonal matrices, which implies that it is block pentadiagonal. In fact, further properties of spherical harmonics show that this non-uniform block pentadiagonal matrix is nonzero only when $l = l' \pm 2$. Hence, the even and odd moments decouple in the second-order term. Figure 1 illustrates this structure for $l = 4$, where the '$\times$' blocks are the non-zero block entries and the small diagonal blocks are the $lm - lm$ 3×3 tensors. Finer structure of this block pentadiagonal matrix can be found by using additional properties of the spherical harmonics (with respect to m).

Now, assume an h finite element discretization for the spatial component. Using the test function $b_\beta(\mathbf{x})Y_{l'm'}(\Omega)$, where $\{b_\beta\}_\beta$ is a basis set for the spatial finite element space, the second-order term becomes

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{lm}\sum_{\alpha} < Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} > (b_{\alpha,i}, b_{\beta,j})_R \, \phi_{lm,\alpha} =$$

$$\sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{lm}\sum_{\alpha}\left[\sum_{l''m''} c_{l''m''} < Y_{lm}|\Omega_i|Y_{l''m''} >< Y_{l''m''}|\Omega_j|Y_{l'm'} >\right]$$

$$(b_{\alpha,i}, b_{\beta,j})_R \, \phi_{lm,\alpha}. \qquad (8)$$

Here, if b_α at spatial node α is the standard hat function, then $\phi_{lm,\alpha}$ is the value of the lm moment at that node. Using the structure of $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$, the structure of the full discretized second-order term is also block pentadiagonal with a total size of $M(N+1)^2$, where M is the total number of spatial nodes. Corresponding to each 3×3 tensor of $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$ is an

$M \times M$ submatrix describing the discretized spatial coupling of moment lm to $l'm'$. Alternatively, assuming R to be decomposed into cubic element and assuming $\{b_\alpha\}$ to be trilinear finite elements, the second-order term can be re-ordered to have a 27 block stripe structure corresponding to the 27 point stencil of the spatial differentiation operator. Each block on any stripe gives the $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$ coupling at a spatial point. Such an ordering is better for computation, but is harder to visualize. Nevertheless, both orderings give an idea of the complexity of the linear system corresponding to a discretization of the PN equations.

Since bilinear form $b(\cdot,\cdot)$ also contains (6), an effective solver or preconditioner must be able to efficiently invert this complex linear system. However, for some parameter regimes, a more sophisticated spatial discretization and a non-standard multigrid scheme may be needed. A problem arises because scaling coefficients c_1 and c_2 may differ by orders of magnitude. Thus, on the one hand, the scaling operator leads to the correct asymptotic limiting solution in these regimes (c.f., [1]), but on the other hand, a complicated discretization and multigrid scheme may be required.

To see this scaling problem, from (7), we see that only the $l''m'' = 00$ column and row of $< Y_{lm}|\Omega_i|Y_{l''m''} >$ and $< Y_{l''m''}|\Omega_j|Y_{l'm'} >$ respectively are scaled by $\sqrt{c_2}$. All other rows and columns are scaled by $\sqrt{c_1}$. Because $< Y_{lm}|\Omega_i S^{-1}\Omega_j|Y_{l'm'} >$ is nonzero only when $l' = l \pm 2$, then only diagonal tensored moment blocks $ll = 00$ and $ll = 11$ of the continuous second-order term can contain c_2 scaled terms. In particular, for regions 1, 2, and 3 respectively, these tensored blocks are

$$\begin{bmatrix} -\frac{1}{3}\nabla\cdot\nabla & 0 \\ 0 & -\frac{1}{5}\underline{\Delta} - \frac{6}{15}\nabla\nabla\cdot \end{bmatrix} \qquad c_1 = 1, c_2 = 1,$$

$$\begin{bmatrix} -\frac{1}{3\sigma_t}\nabla\cdot\nabla & 0 \\ 0 & -\frac{1}{5\sigma_t}\underline{\Delta} - \left(\frac{1}{3\sigma_a} + \frac{1}{15\sigma_t}\right)\nabla\nabla\cdot \end{bmatrix} \qquad c_1 = \frac{1}{\sigma_t}, c_2 = \frac{1}{\sigma_a},$$

and

$$\begin{bmatrix} -\frac{1}{3\sigma_t}\nabla\cdot\nabla & 0 \\ 0 & -\frac{1}{5\sigma_t}\underline{\Delta} - \left(\frac{\sigma_t}{3} + \frac{1}{15\sigma_t}\right)\nabla\nabla\cdot \end{bmatrix} \qquad c_1 = \frac{1}{\sigma_t}, c_2 = \sigma_t.$$

Block $ll = 00$ is Laplacian, so poses no problems. However, block $ll = 11$ contains the grad-div operator $\nabla\nabla\cdot$, whose nullspace consists of divergence-free functions. In particular, in region 3 and region 2 when $\sigma_t \gg 1$ and $\sigma_a \approx \frac{1}{\sigma_t}$, this latter block is dominated by the grad-div operator. Hence, divergence-free components are "near" the nullspace this moment block, and these components pose problems for both standard discretization and standard iterative solvers.

These divergence-free components still create difficulty when both the zeroth and second-order terms of $b(\cdot,\cdot)$ are considered. Now the diagonal blocks in

regions 2 and 3 are

$$\begin{bmatrix} \sigma_a I - \frac{1}{3\sigma_t}\nabla\cdot\nabla & 0 \\ 0 & \sigma_t I - \frac{1}{5\sigma_t}\underline{\Delta} - \left(\frac{1}{3\sigma_a} + \frac{1}{15\sigma_t}\right)\nabla\nabla\cdot \end{bmatrix}, \tag{9}$$

and

$$\begin{bmatrix} \sigma_a^2\sigma_t I - \frac{1}{3\sigma_t}\nabla\cdot\nabla & 0 \\ 0 & \sigma_t I - \frac{1}{5\sigma_t}\underline{\Delta} - \left(\frac{\sigma_t}{3} + \frac{1}{15\sigma_t}\right)\nabla\nabla\cdot \end{bmatrix}. \tag{10}$$

These divergence-free components are essentially eigenfunctions of the $ll = 11$ block corresponding to eigenvalue σ_t. Error components of this form again will be poorly damped out by standard iterative solvers. Algorithms that ameloriate this problem will not be examined in this paper.

4 A Multigrid Algorithm

The solution procedure involves minimizing $\mathcal{F}(\psi; q, g)$ over an appropriate subspace of V. To accomplish this, a Rayleigh-Ritz finite element method is used for the spatial discretization and a truncated spherical harmonic expansion is used for the angle discretization. For the spatial discretization, let $\mathcal{T}_h$ be a triangulation of domain R into elements of maximal length $h = \max\{\text{diam}(K)\cdot : K \in \mathcal{T}_h\}$, and let V^h be a finite dimensional subspace of V having the approximation property

$$\inf_{v^h \in V^h} \|v - v^h\|_{1,R} \leq Ch\|v\|_{2,R}$$

for all $v \in \left[H^2(R) \times L^2(S^2)\right]$. The PN-h finite element space is then

$$V_N^h := \left\{ v_N^h \in V^h : v_N^h = \sum_{l=0}^{N} \sum_{m=-l}^{l} \phi_{lm}^h(x) Y_{lm}(\Omega) \right\}.$$

The discrete fine grid minimization problem is

- Find $\psi_N^h \in V_N^h$ such that

$$\mathcal{F}(\psi_N^h; q, g) = \min_{v_N^h \in V_N^h} \mathcal{F}(v_N^h; q, g).$$

Equivalently, the discrete problem is

- Find $\psi_N^h \in V_N^h$ such that

$$a(\psi_N^h, v_N^h) = \left(q, S^{-1}\mathcal{L}v_N^h\right) + 2\int_{\partial R}\int_{\mathbf{n}\cdot\Omega<0} g\overline{v}_N^h |\mathbf{n}\cdot\Omega|\, d\Omega d\sigma$$

for all $v_N^h \in V_N^h$.

In computation, one solves for the coefficients ϕ^h_{lm}.

A standard projection multilevel scheme for solving either discrete problems is fairly straightforward. Let

$$\mathcal{T}_{2^{m-1}h} \subset \mathcal{T}_{2^{m-2}h} \subset \cdots \subset \mathcal{T}_{2h} \subset \mathcal{T}_h$$

be a conforming sequence of coarsenings of triangulation $\mathcal{T}_h$,

$$V_N^m \subset V_N^{m-1} \subset \cdots \subset V_N^2 \subset V_N^1 := V_N^h$$

a set of nested coarse grid subspaces of V_N^1, the finest subspace, and

$$B^j = \left\{ b^j_{\nu,lm} \right\}$$

a suitable (generally local in space) basis set for V_N^j. (For example, B^j may consists of $b^j_{\nu,lm}$'s that are the product of a level j spacial piecewise linear hat function and the spherical harmonic Y_{lm}.) Given an initial approximation ψ_N^j on level j, the level j relaxation sweep consists of the following cycle

- for each $\nu = 1, 2, ..., M_j$,
  for each lm, $0 \le l \le N$, $-l \le m \le l$,

$$\psi_N^j \leftarrow \psi_N^j + \alpha b^j_{\nu,lm},$$

  where α is chosen to minimize

$$\mathcal{F}\left(\psi_N^h + \alpha b^j_{\nu,lm}; q, g\right). \tag{11}$$

Since $\mathcal{F}\left(\psi_N^h + \alpha b^j_{\nu,lm}; q, g\right)$ is a quadratic function in α, this local minimization process is simple, and is, in fact, a Gauss-Seidel iteration. Moreover, note that the loops range over the moments first so that all moments are updated at a fixed spatial node before going to the next spatial node. Note also that the search direction need not be a single element of B^j, but can be a larger subset of B^j. In such case, with this subset denoted by $\mathbf{b}^{\mathbf{j}}_{\boldsymbol{\nu}}$, then one needs to find the $\boldsymbol{\alpha}$ that minimizes

$$\mathcal{F}\left(\psi_N^h + \boldsymbol{\alpha}\mathbf{b}^{\mathbf{j}}_{\boldsymbol{\nu}}; q, g\right).$$

A good choice for $\mathbf{b}^{\mathbf{j}}_{\boldsymbol{\nu}}$ is the subset $\left\{b^j_\nu\right\}_{lm}$ consisting of all $b^j_{\nu,lm}$ at node ν. With this choice, this block Gauss-Seidel iteration simultaneously updates all the moments at a node.

Now given a fine grid approximation ψ_N^1 on level 1, then the level 2 coarse grid problem is to find a correction ψ_N^2 such that

$$\mathcal{F}\left(\psi_N^h + \psi_N^2; q, g\right) = \min_{v_N^2 \in V_N^2} \mathcal{F}\left(\psi_N^h + v_N^2; q, g\right).$$

Having obtained this correction, ψ_N^1 is updated according to

$$\psi_N^1 \leftarrow \psi_N^1 + \psi_N^2.$$

Applying this procedure recursively yields a multilevel scheme in the usual way.

There is one modification that can be made to this multigrid scheme to achieve better computational efficiency. Because the material and scaling coefficients can have fine scale properties, to preserve these fine scale properties, the above multigrid procedure may require fine scale computation even on the coarser levels. For a matrix-free implementation of this multigrid algorithm, the amount of fine-scale coarse grid calculations is not nominal. To obtain better efficiency, the coefficients can be homogenized to have coarse scale resolution. Thus, the coarse grid calculations can be performed with coarse scale computation. A simple homogenization procedure is to appropriately average the fine scale coefficients. For example, the material and scaling coefficients can be arithmetically and harmonically averaged, respectively: if c_{lm,μ_j}^1, σ_{t,μ_j}^1, and σ_{a,μ_j}^1 are the coefficients on level 1 in element μ_j and level 2 element γ is the agglomeration

$$\gamma = \cup_{j=1}^r \mu_j,$$

then

$$c_{lm,\gamma}^2 = \frac{r}{\sum_{j=1}^r \frac{1}{c_{lm,\mu_j}^1}} \tag{12}$$

$$\sigma_{t,\gamma}^2 = \frac{1}{r}\sum_{j=1}^r \sigma_{t,\mu_j}^1 \tag{13}$$

$$\sigma_{a,\gamma}^2 = \frac{1}{r}\sum_{j=1}^r \sigma_{a,\mu_j}^1. \tag{14}$$

Note that harmonic averaging is performed only on the scaling coefficients because only these coefficients occur in the diffusion tensors of the second-order term, and harmonic averaging is more appropriate for these coefficients (c.f., Dendy).

However, homogenization leads to a violation of a projection multilevel principle. Using the fine scale coefficients on all levels and so, applying fine-scale coarse grid operations, the least-square functional norm is guaranteed to decrease at each step of this multigrid process. In particular, at each step of the relaxation and coarse grid correction procedures, a subspace minimization correction is performed, which implies a decrease in the least-squares functional norm. But since homogenization changes the coefficients on each level, a coarse grid problem corresponds to a minimization problem with a different least-squares functional. Hence, a coarse grid correction is not a subspace correction to the fine level problem. Indeed, for rapidly varying coefficients, it is possible

for a coarse grid correction from a very coarse level to completely pollute the fine level approximation. Thus, extra measures must be taken when applying coefficient homogenization.

Now whether the coefficients are homogenized or not on the coarser grids, by viewing this algorithm from a projection multilevel or subspace correction viewpoint and noticing that all the multigrid components can be performed with spatially local operations, one sees that parallelization is a minor issue. First, one needs to partition the finest computational grid for good processor load balancing. For example, assume that R and the processor topology are rectangular solids with an equal number of processors in each direction. Then since these 3-dimensional domains can be written as a tensor product of three 1-dimensional domains, the processor decomposition can be viewed as a tensor product of three 1-dimensional processor topology decompositions. Each 1-dimensional grid domain is partitioned as equally as possible among the processors in that direction. The tensor product of these three 1-dimensional processor topology decompositions then gives the complete 3-dimensional processor topology decomposition on the finest level.

Next, one needs to introduce a processor topology decomposition on the coarser levels. Unfortunately, as we descend to coarser levels, it is possible to have idle processors when the number of processors exceeds the number of grid nodes. Nevertheless, this multigrid load-balancing bottleneck is not too serious.

Each processor must also have x layers of ghost nodes in each direction. This ensures that each processor has all the information needed to perform a calculation without processor waiting time. This means that appropriate globally updated information must be communicated to other processors. This occurs, for example, after relaxation, residual calculation, and integrid transfer operations. For relaxation, this also implies a multi-colour ordering of the spatial nodes. For a 27-point spatial finite element discretization, an 8-colour ordering is needed. Without this re-ordering, a Gauss-Seidel sweep is not performed, and so, the global least-squares functional need not decrease after each step of this relaxation.

All of the above parallel implementation features were used in our parallel code.

5 A Preconditioned Conjugate Gradient Algorithm

It was shown in [2] that the diagonal $lm - lm$ block preconditioner does not scale with the number of spatial nodes. One reason for this poor performance is the strength of the neglected intra-moment and inter-moment coupling in that block preconditioner. By using a block diagonal preconditioner with the blocks being the full intra-moment coupling, weak scalibility can be achieved. Each of

these intra-moment blocks can be approximately inverted with a few cycles of the above multigrid scheme restricted to that intra-moment block. Comparing the performance of this algorithm to the above multigrid algorithm will expose the relative strength of the intra-moment and inter-moment coupling of the PN equations. This will be numerically analyzed in section 6.

6 Numerical Experiments

The above P_n-h finite element discretization of the FOSLS formulation of the transport equation was implemented. Angle integrals involving spherical harmonics were computed using analytical formulas, and the spatial moments were discretized with piecewise trilinear functions on rectangular solids. Using the above observation of the problematic P1-P1 sub-system in region 3, only regions 1 and 2 were considered- the goal in these experiments was to investigate scalability with increasing number of nodes. We studied 4 cases P_n : n=1, 3, 6, and 9. In all cases, the coefficients are constants. In the test problem, the right hand side is set to zero and a random function is chosen as an initial guess. For a fixed number of moments, we tabulated the the iteration count and the solve time against the number of spatial nodes in Tables 1 and 2. The results for a V-cycle and a W-cycle are recorded in Tables 1 and 2 respectively. In each table, we also record the results for runs with coefficients in regions 1 and 2. In region 1, we found that the V-cycle iteration slows down because the V-cycle does not damp out all the smooth modes of the FOSLS's operator. However these smooth modes were damped out by a multi-grid W-cycle. This is shown by the fact the W-cycle method needs fewer iterations to converge in region 1 than the v-cycle method. The reason is that the W-cycle puts more effort into coarse grid smoothing than a multi-grid V-cycle.

A perusal of the data in the tables reveals that the algorithm scales spatially. We also note that the the iteration count increases linearly with increasing number of moments. While the multigrid W-cycle takes fewer iterations to converge than the multi-grid V-cycle, the multigrid W-cycle takes more time to converge than the multi-grid V-cycle. It is also clear from Table 1 that the multigrid V-cycle is an effective solver for problems in region 2, but is not as good in region 1.

7 Future Research

We have formulated a theory to handle the grad-div operator in region 3. It involve a Helmholtz decomposition of the P_1 moments. Divergence free finite element basis can also be used, but are of higher order and thus more difficult to code than linear elements. We can generalize the code to solve the multigroup BTE. We believe that we can multi-grid the group discretization.

					Region 1		Region 2	
Pn	Nodes	Unk	Proc	Unk/proc	Iter	Time	Iter	Time
1	33^3	4×33^3	1	143,748	9	279	10	313
1	65^3	4×65^3	8	137,313	9	261	10	294
1	129^3	4×129^3	64	134,168	9	267	10	296
1	257^3	4×257^3	512	132,614	9	270	10	302
3	17^3	16×17^3	1	78,608	20	511	14	359
3	33^3	16×33^3	8	71,874	20	472	15	356
3	65^3	16×65^3	64	68,656	20	495	17	413
3	129^3	16×129^3	512	67,084	20	534	19	467
6	9^3	49×9^3	1	35,721	34	1,381	12	496
6	17^3	49×17^3	8	30,092	41	1,493	19	698
6	33^3	49×33^3	64	27,514	43	1,607	24	903
6	65^3	49×65^3	512	26,282	45	1,743	29	1,142
6	17^3	49×17^3	1	240,737	40	9,224	20	4,591
6	129^3	49×129^3	512	205,444	47	10,191	30	6,320
9	9^3	100×9^3	1	72,900	50	8,791	14	2,406
9	17^3	100×17^3	8	61,413	65	9,757	26	3,979
9	33^3	100×33^3	64	56,152	71	10,798	37	5,652
9	65^3	100×65^3	512	53,637	74	11,468	45	7,047

Table 1: V-cycle results for two cases studies of cross sections. In region 1, $\sigma_t = .1$ and $\sigma_a = .05$. In region 2, $\sigma_t = 10$ and $\sigma_a = 5$.

References

[1] T. A. MANTEUFFEL, K. J. RESSEL, AND G. STARKE, *A boundary functional for the least-squares finite element solution of neutron transport problems*, SIAM J. Numer. Anal., 37 (2000), pp.556-586.

[2] B. CHANG AND B. LEE, *Space-Angle First-Order System Least-Squares (FOSLS) for the Linear Boltzmann Equation*, preprint.

					Region 1		Region 2	
Pn	Nodes	Unk	Proc	Unk/proc	Iter	Time	Iter	Time
1	33^3	4×33^3	1	143,748	9	347	10	381
1	65^3	4×65^3	8	137,313	9	337	10	370
1	129^3	4×129^3	64	134,168	8	329	10	404
1	257^3	4×257^3	512	132,614	8	446	10	625
3	17^3	16×17^3	1	78,608	17	541	14	446
3	33^3	16×33^3	8	71,874	17	522	16	497
3	65^3	16×65^3	64	68,656	17	576	17	582
3	129^3	16×129^3	512	67,084	17	722	18	821
6	9^3	49×9^3	1	35,721	33	1,589	12	599
6	17^3	49×17^3	8	30,092	34	1,618	19	918
6	33^3	49×33^3	64	27,514	35	1,865	24	1,327
6	65^3	49×65^3	512	26,282	36	2,656	28	1,961
6	17^3	49×17^3	1	240,737	34	9,662	20	5,749
6	129^3	49×129^3	512	205,444	34	10,487	30	9,537
9	9^3	100×9^3	1	72,900	50	10,017	14	2,860
9	17^3	100×17^3	8	61,413	57	11,086	26	5,235
9	33^3	100×33^3	64	56,152	59	12,561	38	8,264
9	65^3	100×65^3	512	53,637	60	15,548	45	12,017

Table 2: W-cycle results for the same case studies in Table 1.